

# A Simple Nomogram Developed for Predicting HCC Metastasis Based on Micrnas

**Yong Zhu**

Anhui Medical University

**Yusheng Jie**

Third Affiliated Hospital of Sun Yat-Sen University

**Yuankai Wu**

Third Affiliated Hospital of Sun Yat-Sen University

**Wenting Tang**

Sun Yat-sen University Cancer Center

**Jing Cao**

Third Affiliated Hospital of Sun Yat-Sen University

**Zhongzhen Su**

Fifth Affiliated Hospital of Sun Yat-sen University

**ZhenJian Zhuo**

Guangzhou Medical University

**Jiao Gong**

Third Affiliated Hospital of Sun Yat-Sen University

**Yutian Chong** (✉ [chongyt@mail.sysu.edu.cn](mailto:chongyt@mail.sysu.edu.cn))

Third Affiliated Hospital of Sun Yat-Sen University <https://orcid.org/0000-0001-8215-4393>

---

## Primary research

**Keywords:** HCC metastasis, Nomogram, miRNA signature, bioinformatic analysis

**Posted Date:** April 13th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-398595/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

## Background

Owing to lack of predictive models for HCC metastasis based on the expression of miRNAs, we aimed to develop a simple model for identification of HCC patients at high risk of metastasis.

## Methods

HCC datasets with metastasis information were acquired from the Gene Expression Omnibus (GEO), and samples were randomly divided into training group (n=169) and testing group (n=72). Based on expression of miRNAs in the training group, we developed a predictive nomogram for HCC metastasis and evaluated its performance using area under the receiver operating characteristic curve (AUC), calibration curve, decision curve and clinical impact curve analysis.

## Results

We found that the expression of miR-30c, miR-185 and miR-323 in HCC correlated with metastasis by the least absolute shrinkage and selection operator regression (LASSO) method and multivariate logistic regression. Based on these three miRNAs, we generated the nomogram for predicting metastasis in the training group (AUC 0.869 [95% CI .813-0.925], sensitivity 80.5%, specificity 78.9%); in testing group (0.821 [0.770-0.872], 48.5%, 92.3%). The calibration curve showed a good agreement between actual observation and prediction by nomogram. The nomogram represented high clinical net benefits using decision curve and clinical impact curve analysis. Moreover, total scores calculated by nomogram were higher in dead patients than that in alive patients. In addition, the predicted target genes of these 3 miRNAs correlated with tumor metastasis by functional enrichment analysis, such as filopodium. Our easy-to-use nomogram could assist in identifying HCC patients at high risk of metastasis, which offer valuable information for clinical treatment.

## Introduction

Hepatocellular carcinoma (HCC) with several characteristics— rapid progression, poor prognosis, high invasive and resistance to anti-cancer treatments— is a leading cause of cancer-related death worldwide[1, 2]. Although considerable advances in HCC treatment have been achieved, such as resection, ablation and immune therapy, the grim outcome are mainly ascribed to intrahepatic metastasis or recurrence. Although surgical resection is referred to as a potentially curative treatment, almost 70% of HCC patients eventually presents recurrence after resection, and postoperative 5-year survival is less than 30%-40%[3, 4].

microRNAs (miRNAs), small and noncoding RNA gene products, aberrantly expressed in HCC compared with normal tissues, and detecting their expression to classify cancer subtype and predict prognosis is an alternative strategy[5]. Abnormal miRNA expression was involved in metastasis of HCC[6]. For instance,

miR-25 overexpression could promote invasion, epithelial-mesenchymal transition (EMT) formation by targeting Rho GDP dissociation inhibitor alpha (RhoGDI1) in HCC; silencing miR-345 and miR-638 would enhance invasion and EMT of HCC cells[7].

So far, most research mainly focused on the prognostic role of single coding or noncoding gene, several coding genes (mRNA signature) in HCC[8–12]: TXNDC12 could activate  $\beta$ -catenin, and thus trigger EMT and metastasis of HCC cells via activation of; FOXM1[13]; a six-gene metastasis signature composed of 5 mRNAs and 1 lncRNA, or a three-gene signature composed of 3 mRNAs, could successfully predict the probability of metastasis[14]; a 17 gene signature composed of immune cytokines was a superior predictor of venous metastases[10].

To date, miRNA signature for predicting HCC metastasis was limited. One study developed 20-miRNA signature to predict metastasis in HCC patients[4]. Predictably, it is not cost-effective to detect more than 15 miRNAs. Therefore, we sought to construct a simple miRNA signature to accurately predict venous metastasis in HCC patients, which could provide some information for cancer treatment in clinical management. In this study, we identified three independent predictive factors—miR-30c, miR-185 and miR-323—based on these 3 miRNAs, we developed a good nomogram with high specificity and sensitivity for evaluating risk of metastasis. In addition, we predicted targets of these 3 miRNAs, and the target genes correlated with tumor metastasis by functional enrichment analysis, such as filopodium.

## Materials And Methods

### Metastasis related miRNAs expression levels and clinical data from GEO

Expression data and clinical information of GSE6857 were downloaded from Gene Expression Omnibus (GEO, <http://www.ncbi.nlm.nih.gov/geo/>). To assure the accuracy of results, each miRNA with more than 5 missing values was deleted for further analysis. For microarray datasets, we filter out eligible records according to the following principles: (1) HCC patients were pathologically diagnosed, and some of them with metastasis; (2) the sample size was more than 30.

### KM-Plot

“Kaplan-Meier plotter” (KM-Plot, <http://kmplot.com/analysis/>) was used to determine the prognostic role of miRNAs[15]. The related miRNAs data were based on CapitalBio miRNA Array. In our study, we analyzed the correlation between the overall survival time of HCC patients and the expression of two important miRNAs (miR-30c and miR-185).

### MiRNA target prediction

The R package *multiMiR*[16] is a comprehensive tool to predict and validate miRNA-target interactions and their associations with diseases and drugs, which utilize external resources, including validated microRNA-target databases (miRecords, miRTarBase and TarBase), predicted microRNA-target databases (DIANA-microT, EIMMo, MicroCosm, miRanda, miRDB, PicTar, PITA and TargetScan) and microRNA-

disease/drug databases (miR2Disease, Pharmaco-miR VerSe and PhenomiR). We used this tool to predict miRNA targets.

## Functional enrichment analyses and Protein-Protein Interaction network analysis

As previously described, R package “clusterProfiler” was used to perform functional enrichment analyses, and Protein-Protein Interaction (PPI) networks were visualized by Cytoscape[17]. Cytohubba plug-in in Cytoscape was used to analyze hub genes by calculating degrees of connectivity. The hub genes are important nodes with many interactions.

## Statistical Analysis

To evaluate the relative importance of each metastasis-related miRNA, we used the least absolute shrinkage and selection operator (LASSO) regression method and multivariate logistic regression.

As previously described, Nomogram was constructed with rms package and its performance was determined by area under the receiver operating characteristic curve (AUC) and calibration (calibration plots) in R[18]. All the statistical analyses were done using R (version4.0.1) other than indicated. Of all miRNAs in the GSE6857 dataset, missing values less than 0.05% of the fields were imputed by expectation-maximization method via SPSS statistical software, version 25 (SPSS, Inc., USA). A *p* value less than 0.05 was considered statistically significant.

# Results

## Identification of metastasis related miRNAs in hepatoma carcinoma

To construct a model to evaluate the risk of metastasis, we first obtained microarray data retrieved from Gene Expression Omnibus (GEO) database, and 241 samples (GSE=6578) were randomly divided into training group (n=169) and testing group (n=72). The Flow chart is illustrated in Figure 1. In the training group, 41 patients were diagnosed with HCC metastasis, while in the testing group, 19 were diagnosed. There was no big difference in metastasis between these two groups ( $p > 0.05$ ). Then we used LASSO regression analysis to screen miRNAs highly associated with metastasis. The results of the training group with maximal AUC showed 8 potential metastasis related miRNAs, including miR-30c, miR-124a, miR-125b, miR-185, miR-206, miR-323, miR-325 and miR-1 (Figure 2A and 1B). After multivariate logistic regression analysis, only miR-30c, miR-185 and miR-323 were independent risk factors for HCC metastasis (Table 1).

## Nomogram construction of hepatoma carcinoma

We developed a nomogram for predicting HCC metastasis, based on 3 independent risk factors (Figure 2C). In the training group, the nomogram performed well to distinguish HCC patients with metastasis from non-metastasis with a high AUC of 0.869 (95% CI 0.813-0.925); the sensitivity and specificity were as high as 80.49% and 78.91%, respectively (Figure 3A). The optimal cutpoint of 131.59 was calculated

by Cutpoint R package, and the patients in testing group were stratified into the high group (total points >131.59) and the low group (total points  $\leq$ 131.59). Consequently, the AUC in the testing group was 0.821 (95%CI 0.770-0.872) for predicting metastasis with a sensitivity of 48.5% and specificity of 92.3% (Figure 3B, Table 2). Moreover, the calibration plot for metastasis probability confer a good agreement between the prediction by our nomogram and actual observation in the training group and testing group, respectively (Figure 3C and 2D).

We further determined clinical applicability of nomogram using decision curve analysis (DCA) and clinical impact curve analysis (CICA). The results suggested our nomogram could have an optimal clinical net benefit (Figure 3E and 2F).

### **Association between nomogram and survival outcome**

To investigate the correlation between nomogram and survival outcome, HCC patients without follow-up were removed from further analysis in the datasets. We calculated the total points of 211 cases (Alive group, n= 152; dead group, n= 59), and observed that the dead group had higher total points compared with alive group (Figure 4A). Also, the prognostic value of miR-30c, miR-185, miR-323 were obtained using KM plotter (Figure 4B). Kaplan Meier overall survival (OS) curves displayed these results: patients with high miR-30c level had a good prognosis compared with cases with low miR-30c level (n=76, HR= 0.55,  $p=0.042$ ); no obvious difference was observed between low and high miR-185 levels; unfortunately, missing expression data for miR-323 in the database resulted in an unknown prognostic role.

### **Functional enrichment analysis of miRNA targets**

We used the *multiMiR* package to predict miRNA-target interactions from 8 external databases, including DIANA-microT, EIMMo, MicroCosm, miRanda, miRDB, PicTar, PITA and TargetScan. For each miRNA, if the miRNA-target interactions were predicted in 3 different database, then these targets were screened out. We then performed functional enrichment analysis. These genes were enrich in the terms of transcription coregulator activity, phosphatidylinositol binding, transcription corepressor activity, kinase regulator activity by Molecular Function (MF) enrichment analysis ; in the terms of glutamatergic synapse, organelle subcompartment, filopodium by Cellular component (CC) enrichment analysis; in the terms of cell-cell adhesion via plasma-membrane adhesion molecules, embryonic organ development by Biological Process (BP) enrichment analysis (Figure 5A-C). These genes were enrich in the terms of Transcriptional Regulation by MECP2, trans-Golgi Network Vesicle Budding by Reactome pathway analysis (Figure 5D).

The PPI network of these miRNA targets is constructed by string (Supplementary Figure 1) , and MCODE (Molecular COmplex Detection ) in Cytoscope was used to generate 3 stable subnetworks (Figure 6A-C). Hub genes for subnetwork 1 included "CBLB, NEDD4, CUCL2", subnetwork 2 included "PICALM, VAMP2 and NECAP1" and subnetwork 3 included "ACTR1A, DYNLL2 and RAB7A". The prognostic value of NEDD4, CUCL2, PICALM, VAMP2, NECAP1, DYNLL2 and RAB7A in HCC were observed except for CBLB and ACTR1A using KM plotter (Figure 6D).

## Discussion

Identification of HCC metastasis will help clinician to make a better decision for HCC patients. In this study, we identified that miR-30c, miR-185, miR-323 were highly associated with metastasis. Based on these three miRNAs, we constructed a simple, useful and practical nomogram to distinguish metastasis from non-metastasis with high sensitivity and specificity; also, our nomogram possessed greatly clinical net benefit using decision curve analysis and clinical impact curve analysis. The scores of nomogram were high in dead HCC patients than that in alive patients, and patients with high miR-30c level tended to have a better prognosis.

Budhu's colleges had identified some metastasis-related miRNAs, including high upregulated miR-219-1, miR-207, miR-185 and miR-338 as well as down-regulated miR-34a, miR-30c and miR-148a[4]. Besides miR-30c and miR-185, we also found a new metastasis-related miRNA—miR-323[4]. It has been reported that both plasma and tissue miR-323 levels were significantly elevated in papillary thyroid cancer patients with metastasis compared with those without metastasis. MiR-323 attenuated lung cancer cell apoptosis by targeting transmembrane protein with egf-like and 2 follistatin domain [19, 20]. Interestingly, on the contrary, miRNA-323 inhibited cell invasion and metastasis in pancreatic ductal adenocarcinoma by direct suppression of SMAD2 and SMAD3[21]. The reasons for the opposite role of miR-323 in metastasis might be dependent on cell context. This phenomenon was very common for miRNAs: cell-type /tissue-specific competitive mRNAs or untranslated region-binding cofactors may have effect on the binding of miR-323 to its targets; miRNA-regulated targets in different cell types depends on the relative concentration of target mRNAs; the location of miRNA targets in the cell may alter the repressive function of miRNA. For example, miR-122-mediated-downregulation of CAT-1 was reduced owing to the relocation of CAT-1 mRNA (from cytoplasmic processing bodies (PBs) to polysomes) triggered by the cellular stress[22].

Reportedly, 20-miRNA tumor signature to predict venous metastasis of HCC, while we used only 3 miRNAs to achieve a good discrimination in identifying HCC patients at high risk of developing metastases. Moreover, application of our nomogram reached a good differentiation with high AUC values of 0.869 and 0.821 in the training group and testing group respectively. The sensitivity and specificity were as high as 80.49% and 78.91%, respectively. Although the sensitivity (48.5%) was decreased in the training group but the specificity (92.3%) was increased. Our model was a relatively cost-effective strategy with only very few miRNAs.

So far, most research mainly focused on the prognostic role of single coding or noncoding gene, several coding genes (mRNA signature) in HCC. Compared with mRNA, miRNA may be the optimal molecular indicators as the following reasons: Each miRNA is capable to affect the expression of hundreds of coding genes, and thus can affect every facet of cell, such as apoptosis, proliferation and stress response[23]; meanwhile, mature miRNAs are relatively stable and easy to detect[24]. Therefore, miRNA signature can be greatly helpful in diagnosis of metastasis. In this study, we predicted targets of these 3 miRNAs, and the target genes correlated with the term of tumor metastasis by functional enrichment

analysis, such as filopodium, cell-cell adhesion via plasma-membrane adhesion molecules, and transcriptional Regulation by MECP2. Filopodium, cell-cell adhesion and MECP2 play important role in metastasis[25–27].

There were some inherent limitations in the study. First, our study had a small sample size. We were very eager to enroll more samples to construct our model. However, dataset with sufficient clinical information was very few, and strict exclusion criteria (sample size > 30) aggravated this situation. Second, external validation was lacked owing to the different sequencing platforms. Therefore, more comprehensive studies with large sample size are required to confirm performance of our nomogram in clinical management.

In conclusion, we observed 3 metastasis-related miRNAs, including miR-30c, miR-185 and miRNA—miR-323. Our nomogram is practical, easy-to-use, highly sensitive and specific to identify HCC patients at risk of metastasis. Our findings could assist in making a better decision about HCC treatment to improve prognosis.

## **Declarations**

### **Acknowledgements**

Not applicable.

### **Author's contributions**

Yong Zhu, Yusheng Jie and Yuankai Wu had the idea for this study. Wenting Tang and Jing Cao obtained the data. Yongzhu, Yusheng Jie and Yuankai Wu undertook the statistical analysis. Yutian Chong, Jiao Gong and Yong Zhu provided statistical advice. All authors contributed to interpretation of the results.

### **Availability of data and materials**

The dataset used during the study are available from the corresponding author on a reasonable request.

### **Funding**

This work was supported by a grant from the National key research and development

program (2018YFC1315400); National Natural Science Foundation of China (81773176); Basic and Clinical Cooperative Research Promotion Program of Anhui Medical University (2019xkjT029) Clinical Medicine Discipline Construction Project of Anhui Medical University(2020lcxk032) ; Guangdong Natural Science Foundation (NO. 2018A0303130070) .

### **Availability of data and material**

All data generated or analyzed during this study are included in this published article and its supplementary information files.

### **Ethics approval and consent to participate**

Not applicable.

### **Consent for publication**

Not applicable.

### **Competing interests**

The authors declare that they have no competing interests.

## **References**

1. Yang JD, Hainaut P, Gores GJ, Amadou A, Plymoth A, Roberts LR. A global view of hepatocellular carcinoma: trends, risk, prevention and management. *Nature reviews Gastroenterology hepatology*. 2019;16(10):589–604.
2. Villanueva A, Hernandez-Gea V, Llovet JM. Medical therapies for hepatocellular carcinoma: a critical view of the evidence. *Nature reviews Gastroenterology hepatology*. 2013;10(1):34–42.
3. Tabrizian P, Jibara G, Shrager B, Schwartz M, Roayaie S. Recurrence of hepatocellular cancer after resection: patterns, treatments, and prognosis. *Annals of surgery*. 2015;261(5):947–55.
4. Budhu A, Jia HL, Forgues M, Liu CG, Goldstein D, Lam A, Zanetti KA, Ye QH, Qin LX, Croce CM, et al. Identification of metastasis-related microRNAs in hepatocellular carcinoma. *Hepatology*. 2008;47(3):897–907.
5. Asano N, Matsuzaki J, Ichikawa M, Kawauchi J, Takizawa S, Aoki Y, Sakamoto H, Yoshida A, Kobayashi E, Tanzawa Y. **A serum microRNA classifier for the diagnosis of sarcomas of various histological subtypes**. *Nature communications* 2019, 10(1).
6. Callegari E, Gramantieri L, Domenicali M, D'Abundo L, Sabbioni S, Negrini M. MicroRNAs in liver cancer: a model for investigating pathogenesis and novel therapeutic approaches. *Cell Death Differ*. 2015;22(1):46–57.
7. Xu X, Tao Y, Shan L, Chen R, Jiang H, Qian Z, Cai F, Ma L, Yu Y. The Role of MicroRNAs in Hepatocellular Carcinoma. *J Cancer*. 2018;9(19):3557–69.
8. Song BN, Chu IS. A gene expression signature of FOXM1 predicts the prognosis of hepatocellular carcinoma. *Exp Mol Med*. 2018;50(1):e418.
9. Yuan S, Wang J, Yang Y, Zhang J, Liu H, Xiao J, Xu Q, Huang X, Xiang B, Zhu S, et al. The Prediction of Clinical Outcome in Hepatocellular Carcinoma Based on a Six-Gene Metastasis Signature. *Clinical cancer research: an official journal of the American Association for Cancer Research*. 2017;23(1):289–97.

10. Budhu A, Forgues M, Ye QH, Jia HL, He P, Zanetti KA, Kammula US, Chen Y, Qin LX, Tang ZY, et al. Prediction of venous metastases, recurrence, and prognosis in hepatocellular carcinoma based on a unique immune response signature of the liver microenvironment. *Cancer cell*. 2006;10(2):99–111.
11. Yi B, Tang C, Tao Y, Zhao Z. Definition of a novel vascular invasion-associated multi-gene signature for predicting survival in patients with hepatocellular carcinoma. *Oncology letters*. 2020;19(1):147–58.
12. Li B, Feng W, Luo O, Xu T, Cao Y, Wu H, Yu D, Ding Y. Development and Validation of a Three-gene Prognostic Signature for Patients with Hepatocellular Carcinoma. *Scientific reports*. 2017;7(1):5517.
13. Yuan K, Xie K, Lan T, Xu L, Wu H. **TXNDC12 promotes EMT and metastasis of hepatocellular carcinoma cells via activation of  $\beta$ -catenin.** *Cell death and differentiation* 2019:1–14.
14. Liu G-M, Zeng H-D, Zhang C-Y, Xu J-W. **Identification of a six-gene signature predicting overall survival for hepatocellular carcinoma.** *Cancer Cell International* 2019, 19(1).
15. Nagy A, Lanczky A, Menyhart O, Gyorffy B. Validation of miRNA prognostic power in hepatocellular carcinoma using expression data of independent datasets. *Scientific reports*. 2018;8(1):9227.
16. Ru Y, Kechris KJ, Tabakoff B, Hoffman P, Radcliffe RA, Bowler R, Mahaffey S, Rossi S, Calin GA, Bemis L, et al. The multiMiR R package and database: integration of microRNA-target interactions along with their disease and drug associations. *Nucleic Acids Res*. 2014;42(17):e133.
17. Gong J, Li R, Chen Y, Zhuo Z, Chen S, Cao J, Chong Y, Hu B. **HCC subtypes based on the activity changes of immunologic and hallmark gene sets in tumor and nontumor tissues.** *Briefings in Bioinformatics* 2021.
18. Gong J, Ou J, Qiu X, Jie Y, Chen Y, Yuan L, Cao J, Tan M, Xu W, Zheng F, et al: **A Tool to Early Predict Severe Corona Virus Disease 2019 (COVID-19): A Multicenter Study using the Risk Nomogram in Wuhan and Guangdong, China.** *Clinical Infectious Diseases* 2020.
19. Liu Y, Li L, Liu Z, Yuan Q, Lu X. **Plasma miR-323 as a Biomarker for Screening Papillary Thyroid Cancer From Healthy Controls.** *Frontiers in Medicine* 2020, 7(122).
20. Fan JM, Zheng ZR, Zeng YM, Chen XY. MiR-323-3p Targeting Transmembrane Protein with EGF-Like and 2 Follistatin Domain (TMEFF2) Inhibits Human Lung Cancer A549 Cell Apoptosis by Regulation of AKT and ERK Signaling Pathways. *Medical science monitor: international medical journal of experimental clinical research*. 2020;26:e919454.
21. Wang C, Liu P, Wu H, Cui P, Li Y, Liu Y, Liu Z, Gou S. MicroRNA-323-3p inhibits cell invasion and metastasis in pancreatic ductal adenocarcinoma via direct suppression of SMAD2 and SMAD3. *Oncotarget*. 2016;7(12):14912–24.
22. Bhattacharyya SN, Habermacher R, Martine U, Closs EI, Filipowicz W. Relief of microRNA-mediated translational repression in human cells subjected to stress. *Cell*. 2006;125(6):1111–24.
23. Shivdasani RA. MicroRNAs: regulators of gene expression and cell differentiation. *Blood*. 2006;108(12):3646–53.
24. Ye J, Xu M, Tian X, Cai S, Zeng S. Research advances in the detection of miRNA. *Journal of Pharmaceutical Analysis*. 2019;9(4):217–26.

25. Jacquemet G, Hamidi H, Ivaska J. Filopodia in cell adhesion, 3D migration and cancer cell invasion. *Curr Opin Cell Biol.* 2015;36:23–31.
26. Beri P, Popravko A, Yeoman B, Kumar A, Chen K, Hodzic E, Chiang A, Banisadr A, Placone JK, Carter H, et al. Cell Adhesiveness Serves as a Biophysical Marker for Metastatic Potential. *Cancer Res.* 2020;80(4):901–11.
27. Jiang W, Liang Y-L, Liu Y, Chen Y-Y, Yang S-T, Li B-R, Yu Y-X, Lyu Y, Wang R. MeCP2 inhibits proliferation and migration of breast cancer via suppression of epithelial-mesenchymal transition. *J Cell Mol Med.* 2020;24(14):7959–67.

## Tables

Table 1. Multivariate analysis of the training group

Predictor	OR	95% CI	<i>p</i>
hsa-mir-185	6.18	1.68-22.74	0.006
hsa-mir-30c	0.28	0.14-0.56	0.0002
hsa-mir-323	3.41	1.67-6.98	0.0008

Abbreviations: OR, Odds ratio

Table 2. Performance of nomogram for prediction of metastasis

	Metastasis vs non-metastasis		
	AUC (95% CI)	Sensitivity (%)	Specificity (%)
Traning group (n=169)	0.869 (0.813-0.925)	80.5	78.9
Testing group (n=72)	0.821 (0.770-0.872)	48.5	92.3

## Figures

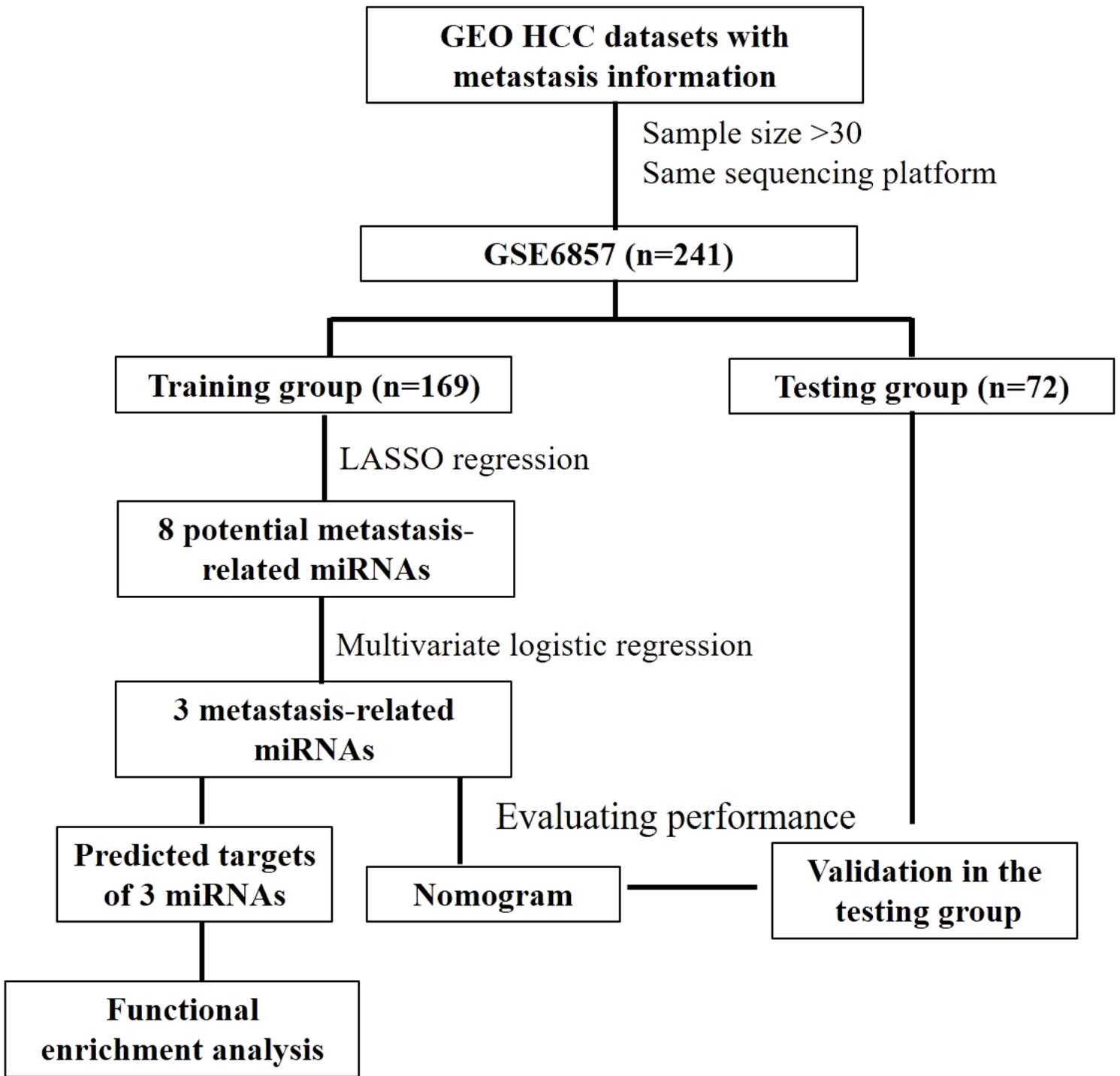
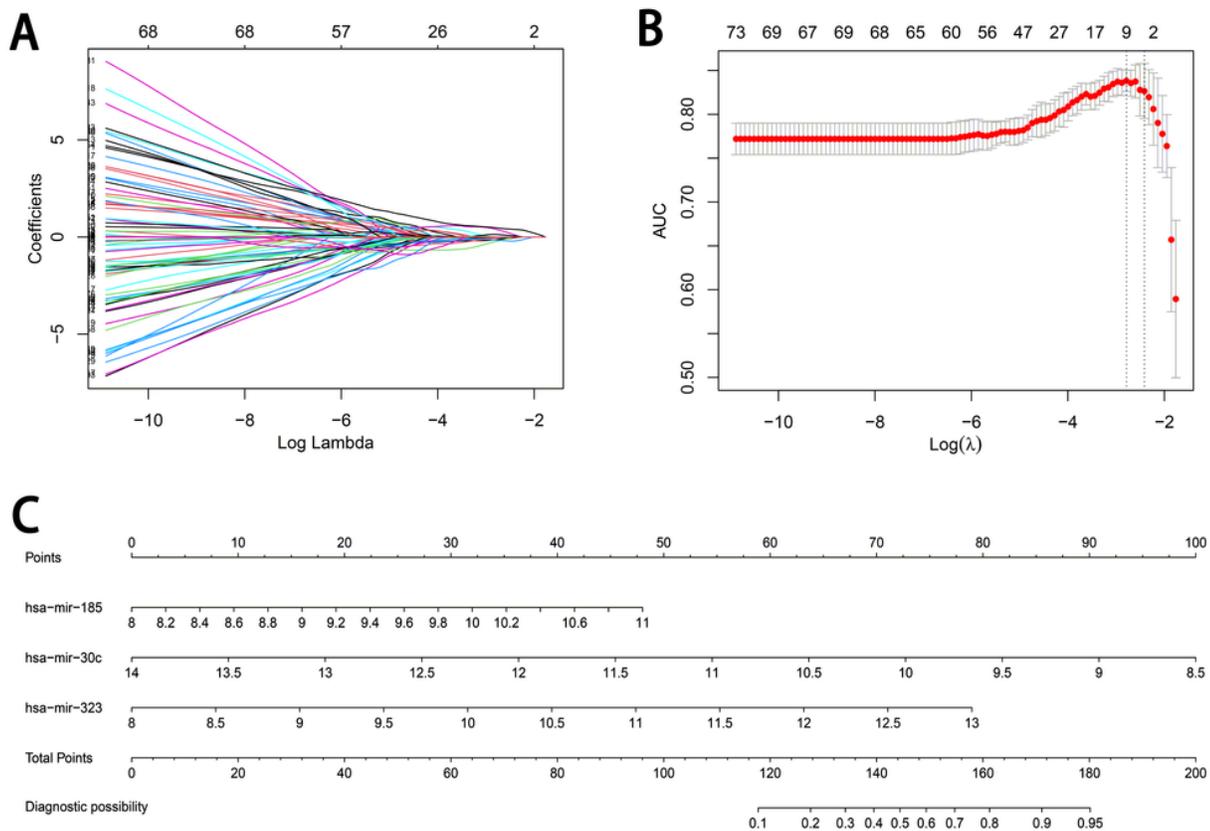


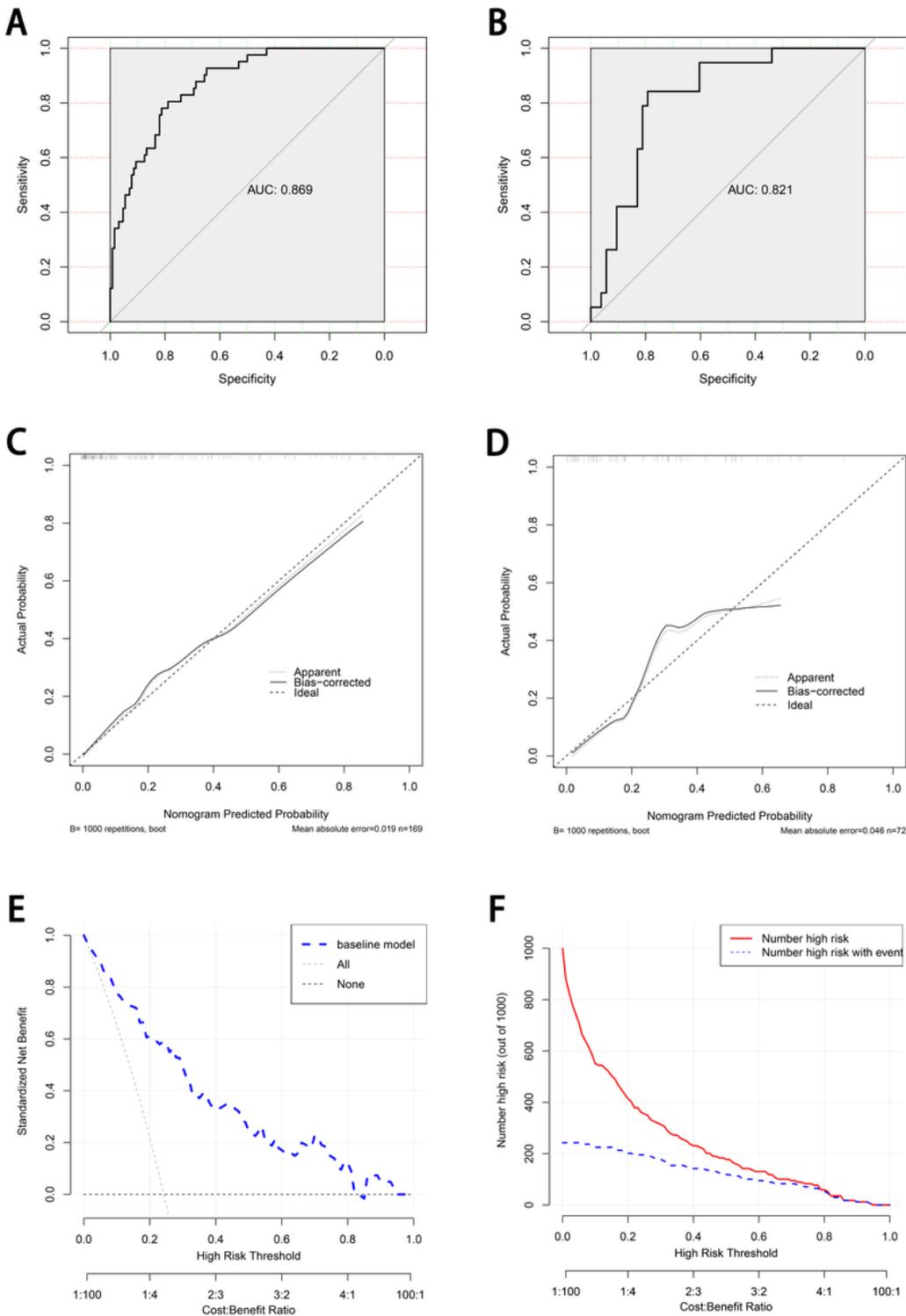
Figure 1

Flow chart of participants in our study.



**Figure 2**

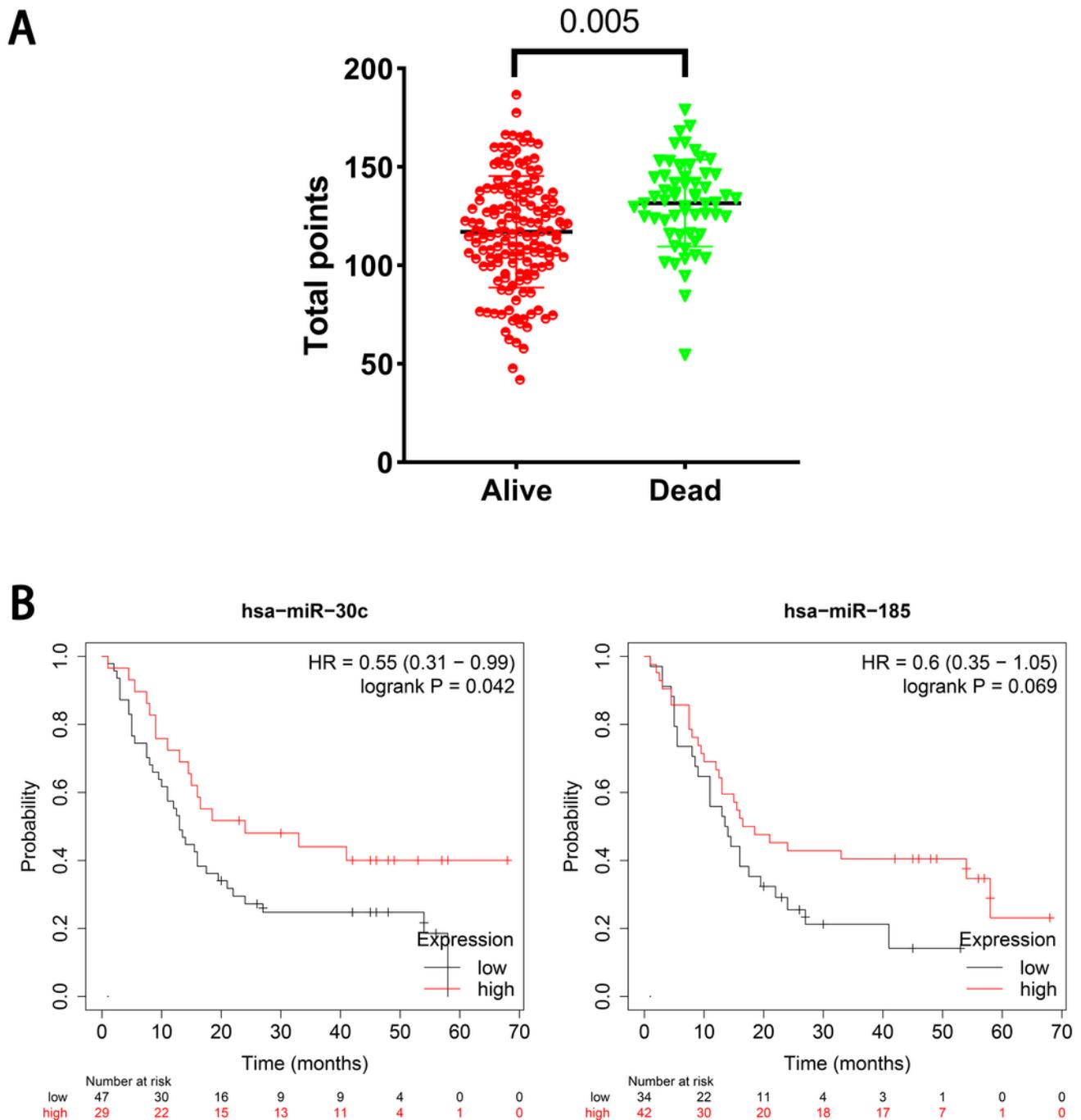
Construction of prediction nomogram in HCC patients with metastasis. The nomogram composed of miR-30c, miR-185 and miR-323 was constructed. (A) LASSO coefficient profiles (y-axis) of the 157 miRNAs. The top x-axis means the average numbers of predictors. The lower x-axis show the log ( $\lambda$ ). (B) Identification of the optimal penalization coefficient ( $\lambda$ ) in the LASSO model with 3-fold cross-validation was based on the maximal area under the receiver operating characteristic (AUC). The y-axis indicates AUC. The dotted vertical lines indicates the optimal values of  $\lambda$ . LASSO, least absolute shrinkage and selection operator. (C) Nomogram predicting the venous metastasis probability in HCC patients was plotted. To read the nomogram in clinical practice, an individual patient's value is laid on each variable axis. We plot an upward line to obtain the number of points for very each miRNA and sum these scores. Then based on the Total points axis, we draw a straight down line to get the probability of metastasis.



**Figure 3**

The calibration and ROC curves of the nomogram The calibration curve, ROC for performance to differentiate HCC patients with metastasis from non-severe metastasis in the training group (A, C) and validation group (B, D), respectively. (E) Decision curve compares the clinical net benefits of three scenarios in predicting the metastasis probability: a perfect prediction model (grey line), screen none (horizontal solid black line), and screen based on the nomogram (blue line). (F) Clinical impact curve

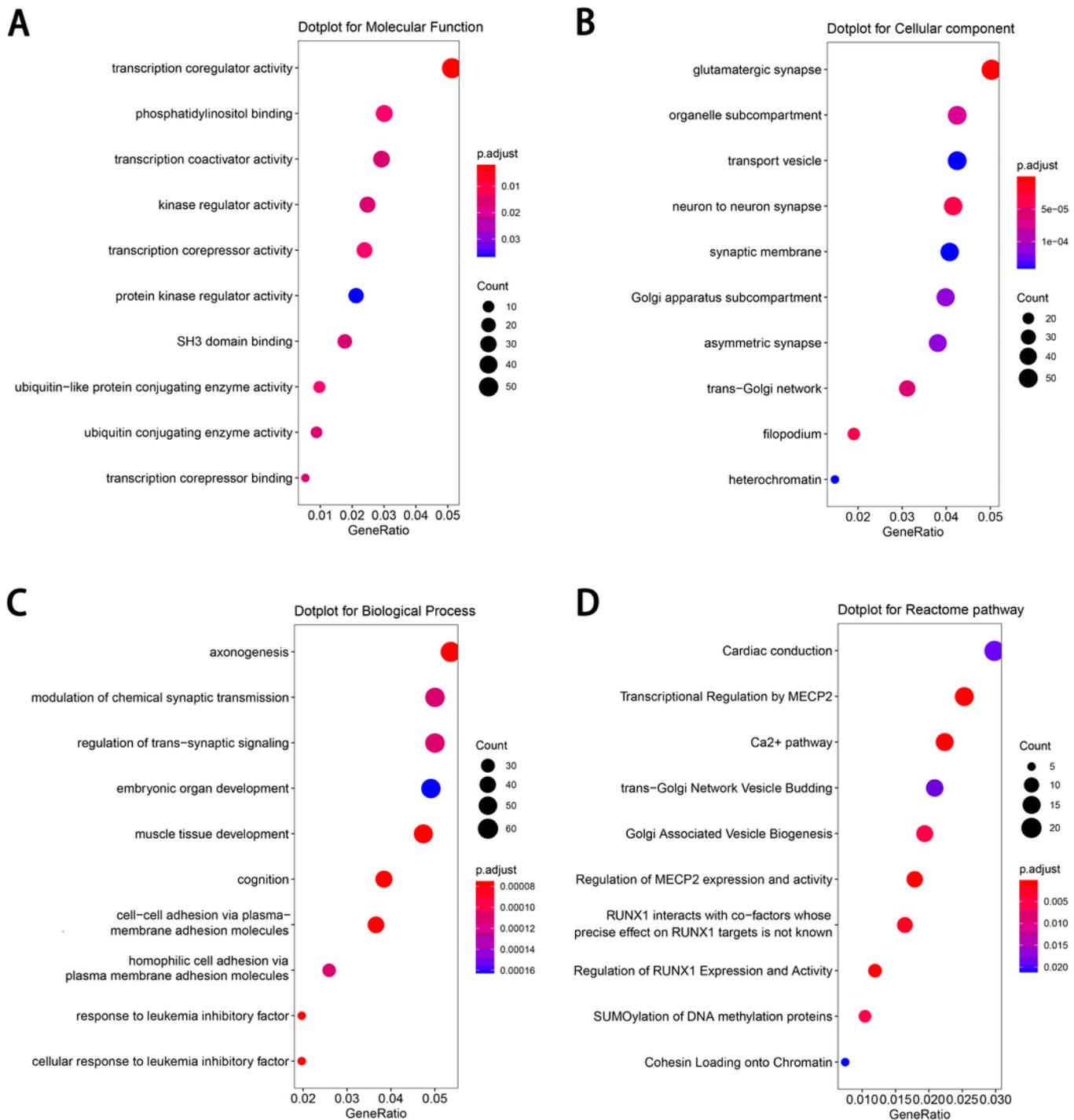
illustrated the number of HCC patients classified as high risk of metastasis, and the number of patients classified high risk with metastasis at each high risk threshold.



**Figure 4**

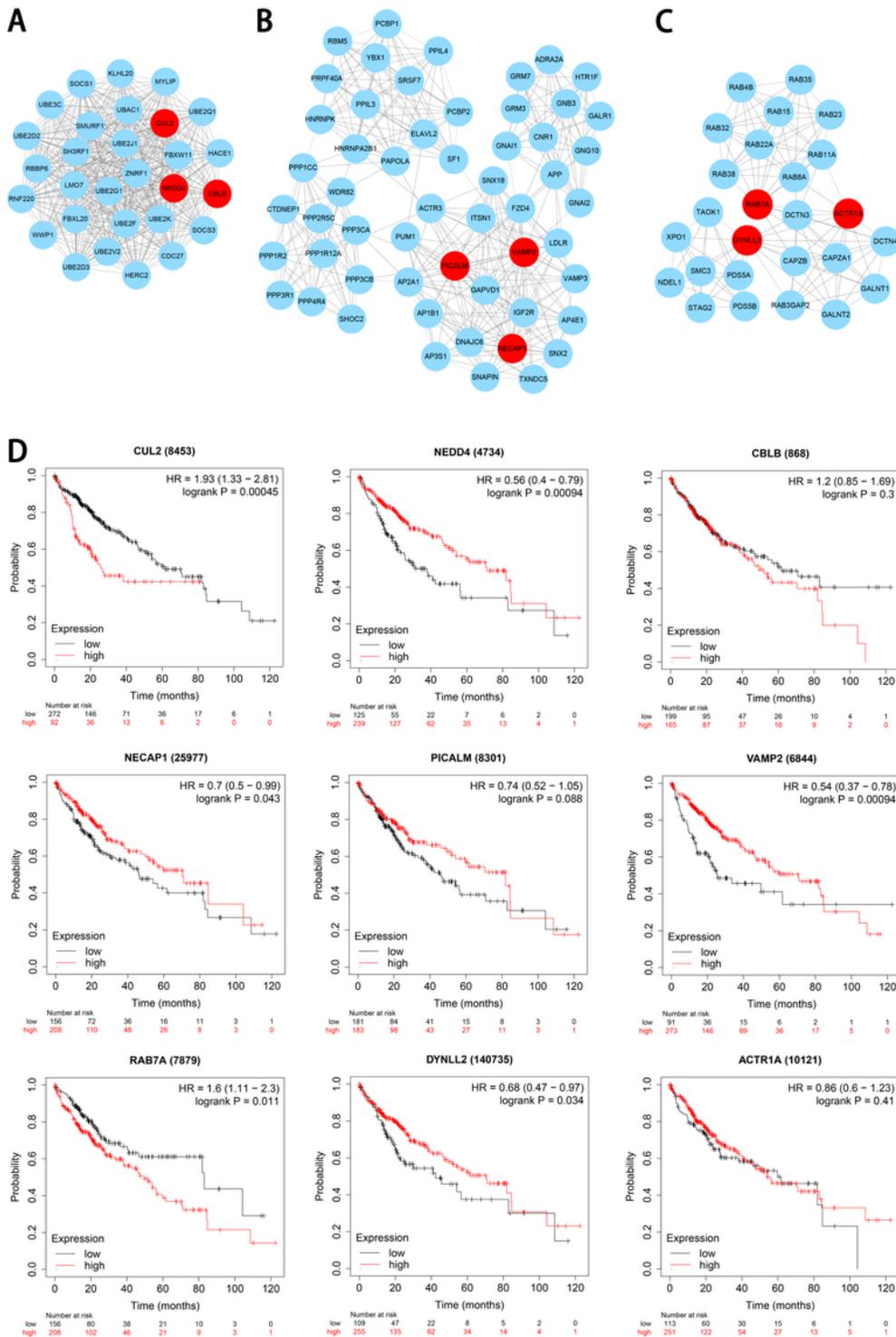
Correlation between the nomogram with survival. (A) Comparison of total points of nomogram between alive group and dead group. (B, C) Kaplan-Meier survival curves analysis was analyzed by KM plotter. The

hazard ratio (HR) and log-rank p-value comparing high and low groups are labeled in chart. HCC patients with high and low miRNAs were indicated in black and red, respectively.



**Figure 5**

Function and Reactome pathway analyses We identified targets of miR-30c, miR-185 and miR-323. These genes subtypes were used to perform Gene oncology (GO) analysis and Reactome pathway analysis.



**Figure 6**

The protein-protein interaction network of miRNA. (A-C) Protein-protein interaction network was generated by string, and was imported into Cytoscape for further analysis with the MCODE plug-in. The top 3 subnetworks were shown. Cytohubba plug-in was used to calculate values of degree, and hub genes (red color) in each subnetworks were identified. (D) Kaplan-Meier survival curves analysis was analyzed by KM plotter.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SFig1.png](#)